



# Identifying four DNA methylation gene sites signature for predicting prognosis of osteosarcoma

Xijun Zhang<sup>1#</sup>, Yongjun Zheng<sup>2#</sup>, Gaoshan Li<sup>3</sup>, Changying Yu<sup>4</sup>, Ting Ji<sup>5</sup>, Shenghu Miao<sup>6</sup>

<sup>1</sup>Department of Laboratory of Jiayuguan City First People's Hospital, Jiayuguan, China; <sup>2</sup>The 984th Hospital of the People's Liberation Army, Shangzhuang Township, Beijing, China; <sup>3</sup>Department of Orthopaedics, 968 Hospital of Joint Service Support Force of Chinese People's Liberation Army, Jinzhou, China; <sup>4</sup>Department of Laboratory Medicine, the 965 Hospital of the PLA, Jilin, China; <sup>5</sup>Shenzhen Mindray Bio-Medical Electronics Co., Ltd, Shenzhen, China; <sup>6</sup>Department of Laboratory Medicine, Wuwei People's Hospital, Wuwei, China

*Contributions:* (I) Conception and design: S Miao, X Zhang; (II) Administrative support: S Wang; (III) Provision of study materials or patients: T Ji; (IV) Collection and assembly of data: G Li; (V) Data analysis and interpretation: Y Zheng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

*Correspondence to:* Shenghu Miao. Department of Laboratory Medicine, Wuwei People's Hospital, Wuwei, China. Email: 381062435@qq.com.

**Background:** Osteosarcoma (OS) is a common malignant bone tumor in children and adolescents. DNA methylation plays a crucial role in the prognosis prediction of cancer. Identification of novel DNA methylation sites biomarkers could be beneficial for the prognosis of OS patients. In this study, we aim to find an efficient methylated site model for predicting survival in OS.

**Methods:** DNA methylation data were downloaded from the Cancer Genome Atlas database (TCGA) and the GEO database. Cox proportional hazard regression and random survival forest algorithm (RSFVH) were applied to identify DNA methylated site signature in the samples randomly assigned to the training subset and the other samples as the test subset. By randomizing 71 clinical samples into two individual groups and a series of statistical analyses between the two groups, a DNA methylation signature is verified.

**Results:** This signature comprises four methylation sites (cg04533248, cg12401425, cg13997435, and cg15075357) associated with the patient training group from the univariate Cox proportional hazards regression analysis, RSFVH, and multivariate Cox regression analysis. Kaplan-Meier survival curves showed the OS patients in the high-risk group have a poor 5-year overall survival compared with the low-risk group, and this finding was identified in the test data set. A ROC analysis was performed in the current research. The results revealed that this signature was an independent predictor of patient survival by investigating the AUC of the four methylation sites signature in the training data set (AUC =0.861) and test data set, respectively (AUC =0.920). The nomogram described in the current study placed a great guiding value for predicting 1-, 2-, 3-year survival of the OS by combining age, gender, grade, and TNM stage as covariates with the RS of patients' methylation related signatures.

**Conclusions:** Our study proved that this signature might be a powerful prognostic tool for survival rate evaluation and guide tailored therapy for OS patients.

**Keywords:** Osteosarcoma (OS); overall survival; prognosis; signature; methylated sites

Submitted Oct 16, 2020. Accepted for publication Nov 18, 2020.

doi: 10.21037/tcr-20-3204

**View this article at:** <http://dx.doi.org/10.21037/tcr-20-3204>

## Introduction

Osteosarcoma (OS) is a common malignant bone tumor in children and adolescents (1). OS incidence worldwide is increasing year by year, accounting for about 35% of primary malignant bone tumors (2). According to the United States' statistical report from 2007 to 2013, OS patients' 5-year relative survival rate was 69.8% from birth to 14 years old and 65.5% from 15 to 19 years old, respectively (3). Despite using multiple chemotherapy regimens before and after surgery, the OS has one of the lowest survival rates of all pediatric cancers (4). Although the development of new adjuvant chemotherapy techniques and surgical methods in recent years has improved OS 5-year survival to 70%, OS mortality and metastasis rates still are high (5,6). Although several anticancer drugs and tumor suppressors have been identified, the underlying molecular mechanisms of OS tumorigenesis remain unclear (7,8). The OS's pathogenesis is not fully elucidated, the critical factor hindering OS treatment progress. Therefore, it is of great significance to clarify the OS's pathogenesis and find the appropriate treatment measures to improve the prognosis and quality of life of patients.

In the past decades, high-throughput technologies, including microarray and gene sequencing, have been widely used to identify driver genes and detect nucleotide polymorphisms and gene fusion in important somatic cells, tumorigenesis, recurrence, and metastasis (9). DNA methylation catalyzed by DNA methyltransferases (DNMTs) is one of the essential epigenetic mechanisms that control cell proliferation, apoptosis, differentiation, cell cycle, and transformation in eukaryotes. Recent progress in epigenetics revealed a deeper understanding of the mechanisms of tumorigenesis and provided biomarkers for early detection, diagnosis, and prognosis in cancer patients (10).

Understanding these genetic changes may shed light on the molecular mechanisms of the OS. However, the inherent genetic and cytogenetic complexity of the OS is challenging because multiple factors mediate tumor biology, including circulating immune cells, hypoxia status, and tumor microenvironment (11). A practical and effective diagnostic test to predict the recurrence risk for OS metastasis or progression is urgently needed. In recent decades, the advent of next-generation sequencing technology has made rapid detection and diagnosis of diseases possible (12). Wang *et al.* suggested patients with high LEVELS of ALDH1B1 expression in OS had lower clinical outcomes than patients with low levels of ALDH1B1, suggesting that the gene may be a potential

prognostic marker in PATIENTS with OS (13). Shi *et al.* studied DDX10 expression differences and prognostic capability for the GEO database. They found the expression level of DDX10 in OS tissues was higher than that in normal tissues, and that increased DDX10 levels were associated with poor prognosis (14).

DNA methylation is an epigenetic change that occurs at the cytosine-phosphate-guanine (CpG) site during the development and progression of various cancers (15). This change has been proven to play a vital role in gene expression, RNA procession, and protein function regulation (16). Some emerging research has recently demonstrated that aberrant DNA methylation is the common and early event in OS, which predates malignant cell proliferation (17). Since DNA methylation genes are strongly associated with tumor metastasis and invasion, the detection of these genes with high accuracy has great significance to the OS early diagnosis (18). It has been reported that tumor-specific methylation sites play a critical role in the early detection and prognosis of cancer, usually overexpressed in cancer cells, while merely expressed in the normal (19). All these findings indicated that these tumor-specific methylation sites have enormous potential for cancer screening. Most studies focused on the relationship between methylation and early diagnosis (20). Up to now, there has been extraordinarily little research reported on using methylation sites as a signature in OS prognosis to investigate the causal relationship between the DNA methylation sites and the clinical outcomes of the OS. Therefore, we aim to find potential survival-related DNA methylation sites signatures in the OS and supply potential therapeutic targets for this, indicating disease. We present the following article in accordance with the TRIPOD reporting checklist (available at <http://dx.doi.org/10.21037/tcr-20-3204>).

## Methods

### *Data retrieval and analysis*

In this reach, the TCGA and GEO database were used to retrieve the clinical and the DNA methylation data of OS patients collected on the Illumina Infinium Human Methylation450 Platform. The methylated information from 237 OS tissue samples was collected from the level 3 methylation databases of the TCGA dataset, and 35 OS tissue samples and seven adjacent normal tissue samples were collected from the GEO database. All these data have been preprocessed. In the TCGA dataset, a total number of 71 samples from OS patients with intact medical records

(gender, age, tumor grade, clinical stage, and vital status) and the methylated information were randomized into two groups: a training group with 47 samples for identifying and constructing the prognostic biomarkers, and a test group with the remaining 24 samples for verifying the accuracy of the prognostic biomarkers. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

### ***Construction of a four-DNA methylation-site signature in the training group***

The clustering analysis was performed between the 35 OS tissue samples and the 7 adjacent normal samples in the GEO dataset. Twenty-one thousand four hundred forty-six methylated differentially expressed gene sites were obtained by considering  $P < 0.05$  and  $FDR < 0.05$  to show a statistically significant difference. The 21,446 methylated differentially expressed gene sites in the 71 OS tissue samples of the TCGA database. A univariate Cox proportional hazard regression analysis was then performed to find the methylation sites related to the patients' OS rate by  $P < 0.05$  in the training group. The random survival forest-variable hunting (RSFVH) algorithm was used to screen for gene sites associated with survival. The Cox evaluated the candidate markers multivariate analysis to screen out the most significant predictive diagnostic and prognostic sites, which were then used to construct the following model to assess the prognosis risk:

$$Risk\ score(RS) = \sum_i^N (expression * coefficient)$$

Here, the *Risk score (RS)* is the methylated differentially expressed DNA methylation site for an OS patient.  $N$  is the representative number of prognostic methylation sites. The *expression* is the expression value of methylation sites, while the *coefficient* is the regression coefficient of methylation sites, the contribution of methylation sites to the prognostic risk score. The patients can be separated into a high-risk and a low-risk group using the median *RS* from the training group as the cutoff point. Subsequently, a risk score system was established, in which the patients with the *RS* higher than the median were arranged into the high-risk group, while the others with lower *RS* were in the low-risk group.

### ***Statistical analysis***

The Kaplan-Meier survival curves were calculated using R software to estimate survival time and compare the high-

and low-risk groups' survival probabilities. Afterward, the time-dependent receiver operating characteristic (ROC) curve was applied to assess the specificity and sensitivity of the methylation sites' survival prediction risk score in the training group. The area under the curve (AUC) was then calculated. The prognostic DNA methylation sites' signatures were constructed by comparing the AUC values in the training group. After that, this methylation signature's prognostic performance was assessed in the test group according to the AUC values and the Kaplan-Meier survival analysis results. Additionally, the evaluated association between the methylation sites expression level and patient survival possibilities was constructed using a nomogram in a training group for the multivariable analysis performed before.

### ***Development of the predictive Nomogram model***

OS patient samples were randomly selected accompanying the intact medical records to develop a predictive nomogram for Iasonos' guideline, which was used in current research to predict the 1-, 2- and 3-year survival rates of OS patients. In brief, the predictive nomogram was developed by combining the *RS* with clinical variables (age, gender, grade, state) using a multivariable Cox regression model. In this model, the Cox proportional hazard multivariate analysis was carried out to identify whether the *RS* for signature methylation sites was an independent survival predictor to assess OS patients' survival rate. All analyses here were conducted with the R program (version 3.5.1) using the following setup: survival random forest (SRC), limma, and pROC packages (Bioconductor, <http://www.bioconductor.org/>). This R program also performed the statistical analysis, and a  $P < 0.05$  was set to determine the statistical significance.

## **Results**

### ***Patient characteristics***

A total of 71 OS patients in the TCGA dataset were used in this study. According to the tumor-node-metastasis (TNM) staging system classification, the tumor's clinical stages are classified into stages I to IV. The clinicopathological characteristics of all 71 OS patients in stages I, II, III, and IV, respectively, were summarized in *Table 1*. The gender, age, grade, stage, and vital status were introduced as variables in this research. If one of these data did not record,

**Table 1** Summary of patient demographics and characteristics

Characteristic	Training (N=47)	Test (N=24)
Gender, n (%)		
Female	23 (48.9)	9 (37.5)
Male	24 (51.1)	15 (62.5)
Age (years), n (%)		
<16	30 (63.82)	20 (83.3)
≥16	17 (36.18)	4 (16.7)
Stage, n (%)		
I	13 (27.7)	7 (29.2)
II	34 (62.3)	17 (70.8)
III	0 (0.0)	0 (0.0)
IV	0 (0.0)	0 (0.0)
Vital status, n (%)		
Living	28 (59.6)	15 (62.5)
Dead	19 (40.4)	9 (37.5)

the related patients were excluded. The cancer cohort samples were randomized into the training and test groups (47 and 24 respectively), and there were 20 and 51 patients in stages I and II in total. The technical route of this study is illustrated in *Figure 1*.

### ***Selection of candidate prognostic methylated sites in the training group***

Clustering analysis was first performed between the OS and the normal tissue samples in the GEO dataset to screen the differentially expressed methylation genes. By setting the thresholds of FDR and  $P < 0.05$ , a total of 21,446 differentially expressed methylated gene sites were picked out. Then called the same differentially expressed methylated gene sites in the TCGA dataset. A univariate Cox proportional hazard regression analysis and RSFVH were then applied to determine the methylated sites in the training group from the dependent variables of survival time and status ( $P < 0.05$ , *Figure 2A,B*). Four methylated sites were teased out, which were the most predictive differentially expressed methylated sites screened out by performing a Cox multivariate analysis. In this work, four methylated sites were identified as strongly correlated to patient survival (cg04533248, cg12401425, cg13997435, and cg15075357) and used to construct a predictive model

to assess the prognosis risk of the OS (*Table S1*).

### ***Building a predictive DNA methylated site signature and confirm of it***

To investigate the performance of the four methylated sites in predicting recurrence, we calculated the *RS* of 4-site signature of all the patients in training set by the *RS* formula as follows:

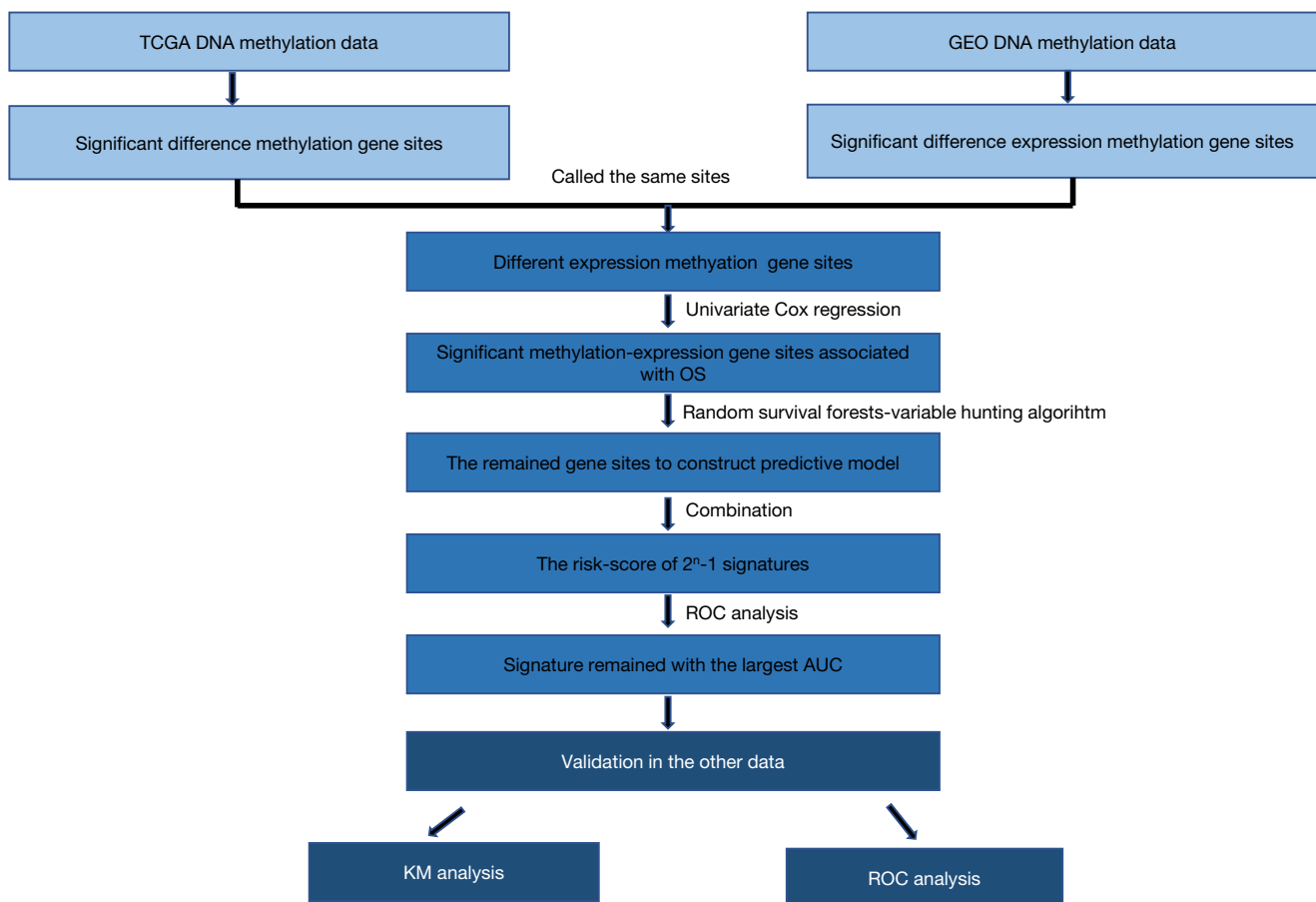
$$RS = (-277.68 \times \text{methylation level value of cg04533248}) + (-372.47 \times \text{methylation level value of cg12401425}) + (11.62 \times \text{methylation level value of cg13997435}) + (3.12 \times \text{methylation level value of cg15075357}).$$

As the median of *RS* in the training set was 0.65, the patients with  $RS > 0.65$  were arranged into the high-risk group, and the others with  $RS \leq 0.65$  were in the low-risk group.

The survival status and distribution of the training set were displayed in *Figure 3*. The results of the Kaplan-Meier survival analysis indicated that the patients in the low-risk group were correlated with a higher survival rate than those in the high-risk group ( $P < 0.001$ ) (OS: 4.4 vs. 2.4 years, log-rank test  $P < 0.05$ ; *Figure 3A*). To confirm the training data set's predicted results, the prognostic *RS*s of 24 patients were calculated in the testing data set in parallel. Following the same logic, these patients were also assigned to the high-risk and low-risk groups, referring to the same *RS* median from the training data set. Patients' survival rate in the high-risk group was significantly lower than patients in the low-risk group (OS: 1.9 vs. 3.5 years, log-rank test  $P < 0.05$ ; *Figure 3B*), which was following the finding from the training set. The *RS* distribution of patients from the training and testing data set was identical.

### ***Building a four methylation sites signature-based predictive nomogram***

To investigate the four-methylated-site signature model's prediction efficacy, the ROC analysis was performed to evaluate the sensitivity and specificity. The AUC of the signature was 0.861 in the training group. This finding indicated that the signature had a good ability to predict OS patients' survival conditions (*Figure 3C*). The conclusion was soon identified in the testing group (AUC = 0.920), which means this signature would hopefully be a powerful prognostic biomarker for OS (*Figure 3D*). We compare the AUC between the signature with other clinical information (gender, age, and stage) to validate its prediction efficacy.



**Figure 1** The flow chart shows the analysis to develop the risk score model in the training group and confirm the model's efficiency in the test group.

We found that the DNA methylation site's signature had the highest AUC value, which indicated the signature had the best sensitivity and accuracy (Figure 3C,D).

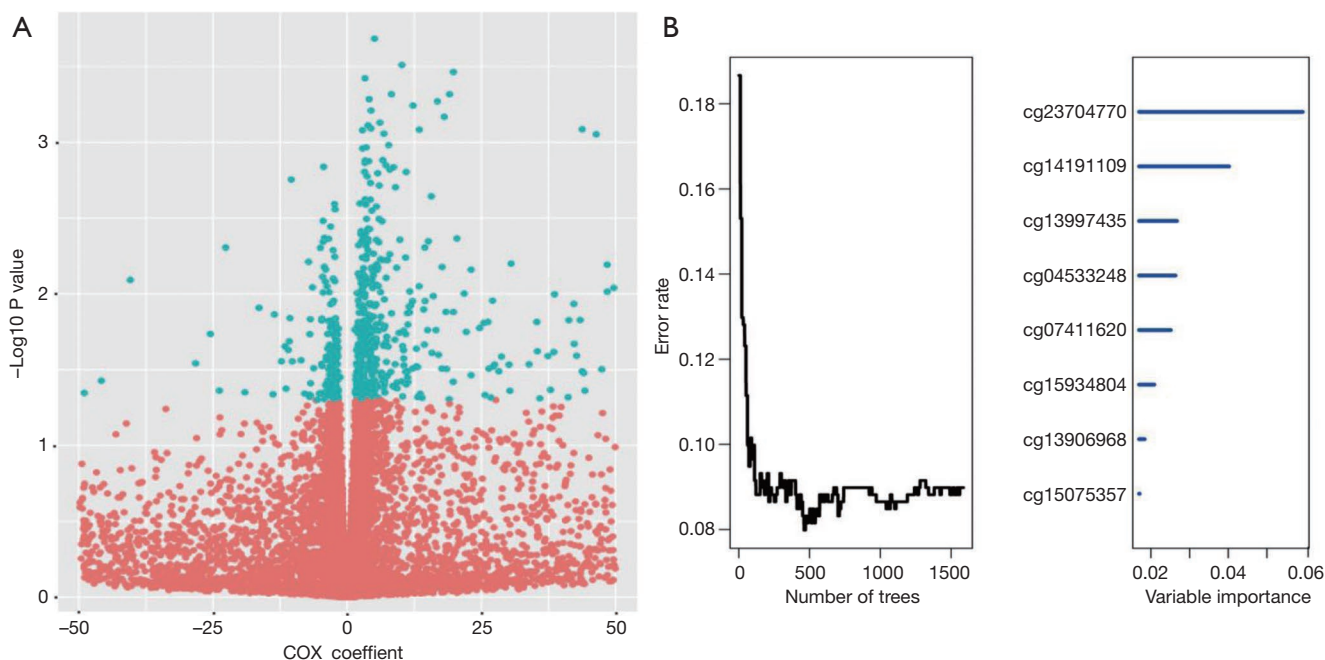
#### ***Building a nomogram consisting of clinical data and RS***

It is widely accepted that the multivariate analysis can reveal the significant association between the RS and the OS. A multivariate Cox regression analysis was performed in 71 patients cohort respectively, in which the age, gender, grade, and stage was set as the co-variables to determine whether the prognostic value of the risk score was an independent variable among the other clinical inputs (age, gender, grade, state). In the current study, it was found the RS (HR =6.26, 95% CI =1.919–20.54,  $P < 0.001$ ) was independently linked with the OS of patients in the multivariate analysis (Figure 4A). To establish a clinically associated effective method that could be used to predict the probabilities of

a 1-, 2-, 3-year OS in OS, a prognostic nomogram from the multivariate analysis results was formulated, and the concordance index of it was 0.85. The predictive nomogram was set up in which the score integrated the other four independent prognostic variables, including age, gender, grade, and stage. The multivariable Cox regression analysis results demonstrated that the prognostic value of the risk score was independent of the four other clinical variables linked with OS patients' overall survival. Further, the prognostic nomogram showed the risk score combined with other clinical data provided a better guiding value for the OS's clinical diagnosis and prognosis than that for the conventional staging systems (Figure 4B).

#### **Discussion**

In this study, we analyzed DNA methylation of OS for understanding the cumulative roles of epigenetic and genetic

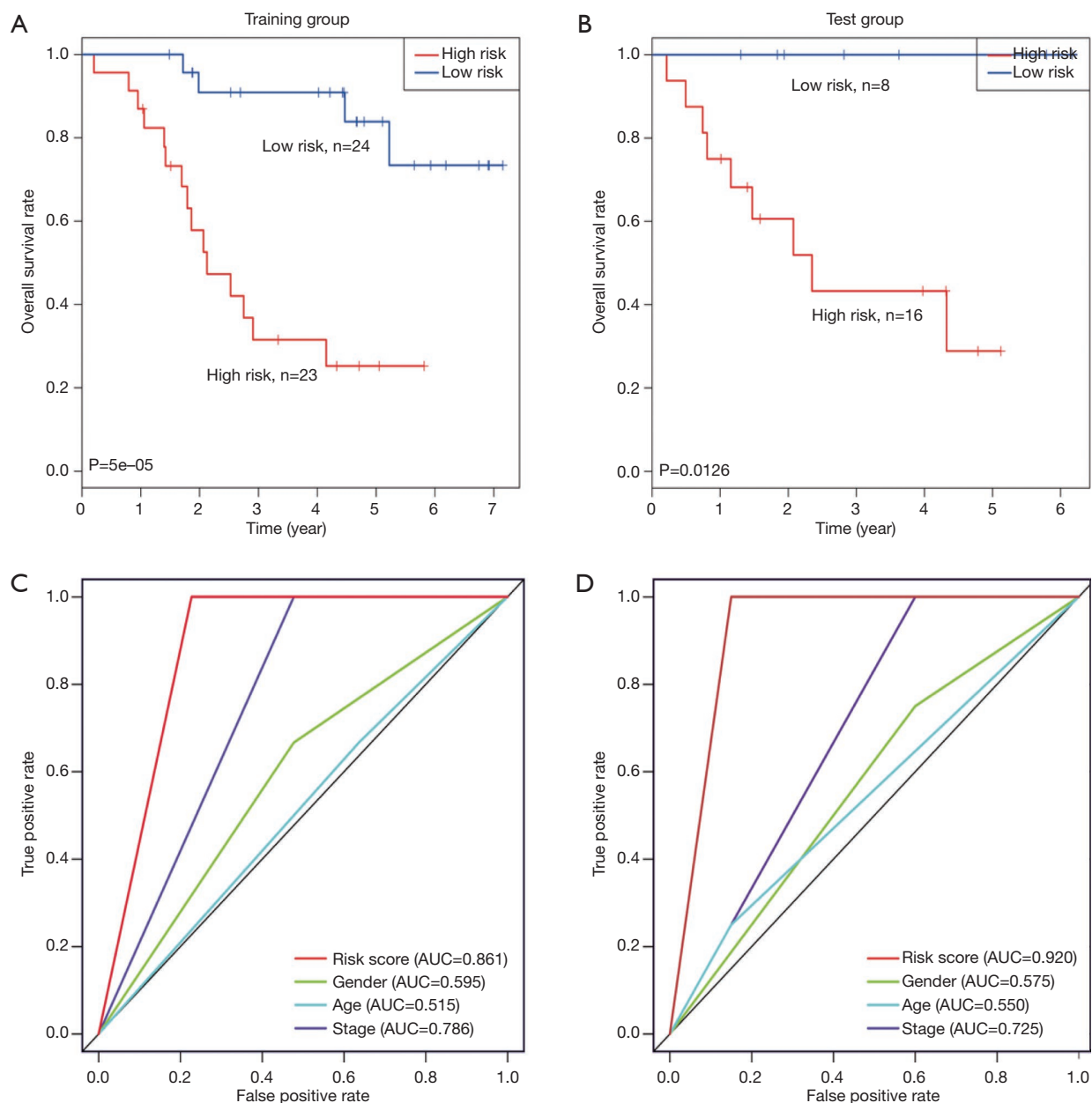


**Figure 2** Screening an optimal model of all methylation gene sites. (A) Cox univariate analysis of the correlation between methylation level of different methylation sites and overall survival, 834 methylation sites were selected by  $P < 0.05$ ; (B) the random survival forests-variable hunting (RSFVH) algorithm screened ten gene sites associated with survival for further study.

mechanisms in OS progression. It is a common cancer with a 5-year survival rate below 50%. The development of a prognosis approach of the OS is urgently needed. The traditional prognostic determinant is with the TNM staging system (21); however, the clinical outcomes are different among patients even at the same stage. This caused the TNM staging system not enough for personalized treatment due to the patients' unpredictable clinical outcomes (22). In the past few years, several powerful predictive biomarkers have been reported (23-25). DNA methylation is a well-known epigenetic change that can alter the expression of vital tumorigenesis associated genes with no genetic sequence changes (26). These changes are strongly associated with the occurrence and development of cancers, and many related biomarkers have been reported (27,28). Tumor-specific methylated sites are critical for the inchoate diagnosis of cancers (29). Since the aberrant DNA methylation was proven to be associated with the OS tumorigenesis, it may serve as a predictive biomarker for OS prognosis (18,30). The DNA methylation genes, including FGD1, METTL3, ALKBH5, and PTEN, are associated with OS tumorigenesis reported in many previous studies (23,31-33) list of them keeps growing. It has been reported that methylation of frizzled-related proteins (SFRPs) may

promote the Wnt signaling pathway, thereby enhancing OS cell invasion (34). Hyper-methylation of p14ADP-ribosylation factor (ARF) and estrogen receptor 1 (ESR1) have been found in OS as well and may have implications in the prognosis of OS patients (35). Moreover, Lu *et al.* have reported that Iroquois homeobox 1 (IRX1) enhances OS metastasis and may be a potential molecular marker (36). A recent study suggests that promoter hyper-methylation of reversion-inducing cysteine-rich protein with Kazal motifs (RECK) is a causative factor in metastasis of OS (37).

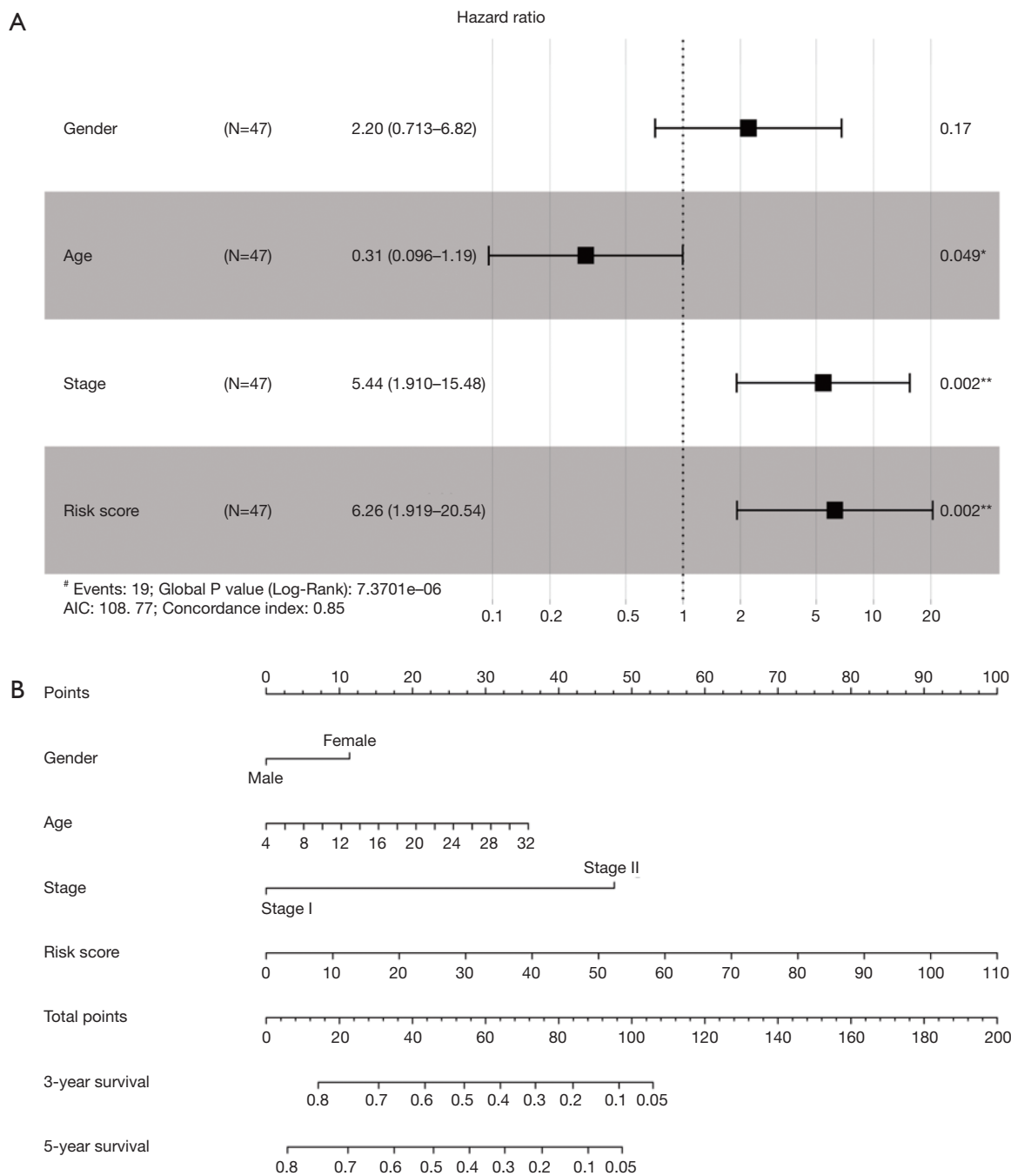
In this study, the methylated genes of OS patients from the TCGA database were carefully collected and analyzed to screen the aberrant methylation DNA related sites, and a novel methylated biomarker model for OS prognostic was successfully built. By randomizing 71 clinical samples into two individual groups and a series of statistical analyses between the two groups, a DNA methylation signature is verified. This signature comprises four methylation sites (cg04533248, cg12401425, cg13997435, and cg1507535) associated with OS patients' survival in the training group, from the univariate Cox proportional hazards regression analysis, RSFVH, and multivariate Cox regression analysis. Kaplan-Meier survival curves and the RS distribution in figure five showed the OS patients in the high-risk



**Figure 3** According to four sifted methylation sites, the patients were stratified into high-risk and low-risk groups from the median of the risk score. Kaplan-Meier survival and ROC analysis were used to visualize the survival probabilities for the two groups. (A) The Kaplan-Meier survival analysis of the OS for TCGA training set patients (n=47); (B) the Kaplan-Meier survival analysis of the OS for TCGA test set patients (n=24); (C) ROC curve for the 5-year survival prediction by combining 4 methylation site signatures in the training dataset; (D) ROC curve for the 5-year survival prediction by combining 4 methylation site signatures in the test dataset. OS, overall survival; TCGA, the Cancer Genome Atlas database; ROC, receiver operating characteristic.

group have a poor 5-year OS compared with the low-risk group, and. This finding was identified in the test data set. As this signature's specificity and sensitivity are critical for OS diagnosis and prognosis, a ROC analysis

was performed in current research. The results revealed that this signature was an independent predictor of patient survival by investigating the AUC of the four methylation sites signature in the training data set (AUC =0.861) and



**Figure 4** Establishment of a nomogram for overall survival prediction in OS. (A) Multivariate analysis revealed a significant association between the risk score and OS (HR =6.26,  $P<0.001$ ); (B) a nomogram combining clinical data (age, gender, grade, and stage) with risk score provides great guiding value for predicting 1-, 2-, 3-year overall OS survival. OS, osteosarcoma. \*,  $P<0.05$ ; \*,  $P<0.01$ ; #,  $P<0.001$ .

test data set, respectively (AUC =0.920). A predictive model with these four methylated sites was then set up to assess each patient’s recurrence probability in the training group by a nomogram. The results showed that this signature

has considerable prognostic power for the OS and could supply robust prediction for each OS patient. In this study, a multivariate Cox regression analysis was used to assess the risk scores supported by an independent correlation with



the OS. Since the TNM staging system alone is insufficient for tumor prognosis as described above, the nomogram described in the current study placed a great guiding value for the predicting of 1-, 2-, the 3-year survival of the OS by combining age, gender, grade, and TNM stage as covariates with the RS of patients with the methylation related signatures.

According to the analysis results, a four-DNA methylation site model was involved in four genes BCAS3, EXO1, S100A2, and NPHP4. Among these genes, BCAS3, EXO1, and S100A2 were already annotated to regulate cancer in earlier research, while NPHP4 was less well-known. It has been reported that BCAS3 is a gene of unknown function on human chromosome 17q23, a region associated with breakpoints of several neoplasms. A study recently reported a conserved protein expressed in the embryonic vasculature and malignant tumors (38). BCAS3 also functions as an interesting protein in breast cancer plasma because their expression level was higher in the breast cancer samples than in healthy donors (39). EXO1 participates in various DNA damage repairs, including mismatch repair, nucleotide excision repair, and homologous recombination (40). A genetic study in yeast shows a role of Exo1 in non-homologous end joining (NHEJ), acting as a regulator for accuracy repairing DNA (41). The S100 calcium-binding protein A2 (S100A2) has been observed in several human cancers (42). S100A2 protein overexpression in CRC cells was associated with significantly worse overall survival and relapse-free survival, showing that S100A2 is an independent risk factor for stage II and III colorectal cancer recurrence (43).

Although some limitations exist in this study, including our study, samples are entirely retrospective, and inherent biases may influence results. And we have not further searched the mechanism action of these DNA methylation genes in OS. Despite these drawbacks, the significant and consistent correlation between our four methylated site signature and overall survival in two independent datasets showed a potentially powerful prognostic marker for the OS. What's more, Detection of methylation technology is relatively mature, can be better used in clinical practice. In summary, we confirmed this signature could serve as potential robust and specificity biomarkers in the prognosis, prediction, and tailored therapy for OS patients.

## Acknowledgments

*Funding:* None.

## Footnote

*Reporting Checklist:* The authors have completed the TRIPOD reporting checklist. Available at <http://dx.doi.org/10.21037/tcr-20-3204>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr-20-3204>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Ritter J, Bielack SS. Osteosarcoma. *Ann Oncol* 2010;21 Suppl 7:vii320-5.
2. Mirabello L, Troisi RJ, Savage SA. Osteosarcoma incidence and survival rates from 1973 to 2004: Data from the Surveillance, Epidemiology, and End Results Program. *Cancer* 2009;115:1531-43.
3. Wakamatsu T, Kakunaga S, Takenaka S, et al. Prognostic implication of adjuvant/neoadjuvant chemotherapy consisting of doxorubicin and ifosfamide in patients with extraskeletal osteosarcoma. *Int J Clin Oncol* 2019;24:1311-9.
4. Bielack S, Carrle D, Jost L. Osteosarcoma: ESMO Clinical Recommendations for diagnosis, treatment and follow-up. *Ann Oncol* 2008;19:ii94-ii96.
5. Ferrari S, Smeland S, Mercuri M, et al. Neoadjuvant Chemotherapy With High-Dose Ifosfamide, High-Dose Methotrexate, Cisplatin, and Doxorubicin for Patients With Localized Osteosarcoma of the Extremity: A Joint

- Study by the Italian and Scandinavian Sarcoma Groups. *J Clin Oncol* 2005;23:8845-52.
6. Yu P, Wen J, Wang J, et al. Establishment and characterization of a novel human osteosarcoma cell line for spontaneous pulmonary metastasis research in vivo. *Ann Transl Med* 2019;7:573.
  7. McQueen P, Ghaffar S, Guo Y, et al. The Wnt signaling pathway: implications for therapy in osteosarcoma. *Expert Rev Anticancer Ther* 2011;11:1223-32.
  8. Wu CL, Tsai HC, Chen ZW, et al. Ras activation mediates WISP-1-induced increases in cell motility and matrix metalloproteinase expression in human osteosarcoma. *Cell Signal* 2013;25:2812-22.
  9. Fox, Samuel E. Transcriptomic analysis using high-throughput sequencing and DNA microarrays. *Dissertations & Theses Gradworks*, 2011.
  10. Pan Y, Liu G, Zhou F, et al. DNA methylation profiles in cancer diagnosis and therapeutics. *Clin Exp Med* 2018;18:1-14.
  11. Zhou J, Wang W, Yan Q, et al. Expression of HER-2 in surgical specimen and biopsy as a biomarker of metastasis in patients with osteosarcoma: a meta-analysis. *Transl Cancer Res* 2019;8:1129-36.
  12. Liu Y, Wei X, Kong X, et al. Correction: Targeted Next-Generation Sequencing for Clinical Diagnosis of 561 Mendelian Diseases. *PLoS One* 2015;10:e0139258.
  13. Wang X, Yu Y, He Y, et al. Upregulation of ALDH1B1 promotes tumor progression in osteosarcoma. *Oncotarget* 2017;9:2502-14.
  14. Shi JH, Hao YJ. DDX10 overexpression predicts worse prognosis in osteosarcoma and its deletion prohibits cell activities modulated by MAPK pathway. *Biochem Biophys Res Commun* 2019;510:525-9.
  15. Saeed H, Sinha S, Mella C, et al. Aberrant epigenetic silencing of neuronatin is a frequent event in human osteosarcoma. *Oncotarget* 2020;11:1876-93.
  16. Yi L, Luo P, Zhang J. Identification of aberrantly methylated differentially expressed genes in breast cancer by integrated bioinformatics analysis. *J Cell Biochem* 2019;120:16229-43.
  17. Asano N, Takeshima H, Yamashita S, et al. Epigenetic reprogramming underlies efficacy of DNA demethylation therapy in osteosarcomas. *Sci Rep* 2019;9:20360.
  18. Manara MC, Valente S, Cristalli C, et al. A Quinoline-Based DNA Methyltransferase Inhibitor as a Possible Adjuvant in Osteosarcoma Therapy. *Mol Cancer Ther* 2018;17:1881-92.
  19. Lorincz AT. Cancer diagnostic classifiers based on quantitative DNA methylation. *Expert Rev Mol Diagn* 2014;14:293-305.
  20. Tian W, Li Y, Zhang J, et al. Combined analysis of DNA methylation and gene expression profiles of osteosarcoma identified several prognosis signatures. *Gene* 2018;650:7-14.
  21. Choi YY, Jang E, Seo WJ, et al. Modification of the TNM Staging System for Stage II/III Gastric Cancer Based on a Prognostic Single Patient Classifier Algorithm. *J Gastric Cancer* 2018;18:142-51.
  22. Jiang W, Liu N, Chen XZ, et al. Genome-wide identification of a methylation gene panel as a prognostic biomarker in nasopharyngeal carcinoma. *Mol Cancer Ther* 2015;14:2864-73.
  23. Chen S, Zhou L, Wang Y. ALKBH5-mediated m6A demethylation of lncRNA PVT1 plays an oncogenic role in osteosarcoma. *Cancer Cell Int* 2020;20:34.
  24. Han X, Liu F, Zhang C, et al. SIAH1/ZEB1/II-6 Axis is Involved in Doxorubicin (Dox) Resistance of Osteosarcoma Cells. *Biol Chem* 2019;400:545-53.
  25. Gong HL, Tao Y, Mao XZ, et al. MicroRNA-29a suppresses the invasion and migration of osteosarcoma cells by regulating the SOCS1/NF- $\kappa$ B signalling pathway through negatively targeting DNMT3B. *Int J Mol Med* 2019;44:1219-32.
  26. Peters FS, Manintveld OC, Betjes MG, et al. Clinical potential of DNA methylation in organ transplantation. *J Heart Lung Transplant* 2016;35:843-50.
  27. Klutstein M, Nejman D, Greenfield R, et al. DNA Methylation in Cancer and Aging. *Cancer Res* 2016;76:3446-50.
  28. Hao X, Luo H, Krawczyk M, et al. DNA methylation markers for diagnosis and prognosis of common cancers. *Proc Natl Acad Sci U S A* 2017;114:7414-9.
  29. Warton K, Mahon KL, Samimi G. Methylated circulating tumor DNA in blood: power in cancer prognosis and response. *Endocr Relat Cancer* 2016;23:R157-71.
  30. Shi X, Fan M. Tip60-dependent acetylation of KDM2B promotes osteosarcoma carcinogenesis. *J Cell Mol Med* 2019;23:6154-63.
  31. Wu W, Jing D, Meng Z, et al. AFGD1 promotes tumor progression and regulates tumor immune response in osteosarcoma via inhibiting PTEN activity. *Theranostics* 2020;10:2859-71.
  32. Ling Z, Chen L, Zhao J. M6A-dependent up-regulation of DRG1 by METTL3 and ELAVL1 promotes growth, migration, and colony formation in osteosarcoma. *Biosci Rep* 2020;40:BSR20200282.
  33. Yao P, Ni Y, Liu C. Long Non-Coding RNA 691

- Regulated PTEN/PI3K/AKT Signaling Pathway in Osteosarcoma Through miRNA-9-5p. *Onco Targets Ther* 2020;13:4597-606.
34. Xiao Q, Yang Y, Zhang X, et al. Enhanced Wnt signaling by methylation-mediated loss of SFRP2 promotes osteosarcoma cell invasion. *Tumour Biol* 2016;37:6315-21.
  35. Hu W, Wu X, Tang J, et al. Anti-cancer targets of formononetin and molecular mechanisms in osteosarcoma: Findings of bioinformatic and experimental assays. *J Cell Mol Med* 2019;23:3505-11.
  36. Lu J, Song G, Tang Q, et al. IRX1 hypomethylation promotes osteosarcoma metastasis via induction of CXCL14/NF- $\kappa$ B signaling. *J Clin Invest* 2015;125:1839-56.
  37. Zhou Z, Wang Z, Wei H, et al. Promotion of tumour proliferation, migration and invasion by miR-92b in targeting RECK in osteosarcoma. *Clin Sci (Lond)* 2016;130:921-30.
  38. Jain M, Bhat GP, Vijayraghavan K, et al. Rudhira/BCAS3 is a cytoskeletal protein that controls Cdc42 activation and directional cell migration during angiogenesis. *Exp Cell Res* 2012;318:753-67.
  39. Corrêa S, Panis C, Binato R, et al. Identifying potential markers in Breast Cancer subtypes using plasma label-free proteomics. *J Proteomics* 2017;151:33-42.
  40. Sousa JF, Serafim RB, Freitas LM, et al. DNA repair genes in astrocytoma tumorigenesis, progression and therapy resistance. *Genet Mol Biol* 2019;43:e20190066.
  41. He D, Li T, Sheng M, et al. Exonuclease 1 (Exo1) Participates in Mammalian Non-Homologous End Joining and Contributes to Drug Resistance in Ovarian Cancer. *Med Sci Monit* 2020;26:e918751.
  42. Wolf S, Haase-Kohn C, Pietzsch J. S100A2 in cancerogenesis: a friend or a foe? *Amino Acids* 2011;41:849-61.
  43. Masuda T, Ishikawa T, Mogushi K, et al. Overexpression of the S100A2 protein as a prognostic marker for patients with stage II and III colorectal cancer. *Int J Oncol* 2016;48:975-82.

**Cite this article as:** Zhang X, Zheng Y, Li G, Yu C, Ji T, Miao S. Identifying four DNA methylation gene sites signature for predicting prognosis of osteosarcoma. *Transl Cancer Res* 2020;9(11):7299-7309. doi: 10.21037/tcr-20-3204